

# 金融实证分析中应注意样本容量的确定

赵 华

**摘 要:**人们在金融实证分析中,很少关注如何确定样本容量(样本大小)的问题。本文以 GARCH 模型为例,通过递归估计参数和相应的矩条件,并结合上海股市的具体情况进行分析,结果表明上海股市波动率 GARCH 模型最优样本容量是三到四年。本文最后从中国金融制度的变化、样本容量对研究结论的影响等方面探讨了样本容量的确定问题。

**关键词:** GARCH 模型;实证金融;样本容量

## Noting the Determination of Sample Size in Empirical Financial Analysis

Zhao Hua

**Abstract:** People rarely pay attention to the determination of sample size in empirical financial analysis. The paper, by GARCH models and through recursive estimation parameters and corresponding moment conditions, empirically analyzes the Shanghai stock market, and reaches the conclusion that the optimized sample size of GARCH model of shanghai stock fluctuation is three to four years. Furthermore, the paper probes into the determination of sample size from the financial system changes of China and the effects of sample size on research conclusion.

**Key words:** GARCH model; empirical finance; sample size

## 1 问题的提出

在投资、证券定价、风险管理和货币政策的实证分析中,样本的选择是第一位的,正确的样本选择是得到正确实证结论的前提条件。我们在教学和研究中发现,人们在实证金融分析中,可能会注意样本数据的平稳性、数据的频率,却很少关注如何确定样本容量。人们往往根据数据的可得性选择样本,基于样本进行实证分析,从而得到结论和相应的政策建议,却没有考虑到样本容量的不同对结论的影响。本文将以前实证金融中使用频率最高的 GARCH 模型为例谈谈样本大小的重要性以及如何确定样本容量。

## 2 GARCH 模型的常规条件

ARCH 模型最初由 Engle (1982) 提出,并由 Bollerslev (1986) 推广形成模型。收益率序列  $\{R_t\}$  的

AR(r)-GARCH(p, q) 模型为

$$R_t = \phi_0 + \phi_1 \sum_{i=1}^r R_{t-i} + \varepsilon_t \quad (1)$$

$$\varepsilon_t = \eta_t \sqrt{h_t}, \quad \eta_t \sim IID(0, 1) \quad (2)$$

$$h_t = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j} \quad (3)$$

这里模型阶数  $r \geq 0$ ,  $p \geq 0$ ,  $q \geq 0$ 。为了保证非负方差,参数  $\omega > 0$ ,  $\alpha_i > 0$ ,  $\beta_j > 0$ 。Bollerslev (1986) 研究表明 GARCH(1, 1) 模型的二阶平稳的充分必要条件是

$$(\alpha + \beta) < 1 \quad (4)$$

Ling 和 McAleer (2002) 得出了 GARCH(1, 1) 四阶矩存在的充分必要条件

$$(\alpha + \beta)^2 + 2\alpha^2 < 1 \quad (5)$$

二阶矩条件和四阶矩条件构成了 GARCH(1, 1) 模型平稳性的常规条件,是得出 GARCH 模型最优样本容量的前提条件。

作者简介:赵华,1975 年生,安徽人。2005 年毕业于厦门大学计划统计系,经济学博士,现为厦门大学金融系副教授,硕士生导师,北京大学中国金融研究中心博士后。

### 3 最优样本容量的确定

#### 3.1 数据描述、模型设定

为了反映上海股市波动率模型与样本容量大小的关系,本文采用上证综指作为代表来描述上海股市的波动,考虑到波动率估计的大样本性质,我们采用日收盘价数据。不同的交易制度下市场的波动会产生显著的不同,不宜跨不同的交易制度选择数据,从 1996 年 12 月 16 日至今,中国股市实行有 10% 涨跌停板限制的 T+1 交易制度,本文数据的时间周期从 1997 年 1 月 2 日至 2004 年 12 月 31 日,共 1928 个交易日数据。对于股票指数收益率的计算,本文采取对数收益率,即

$$shhr_t = \ln shhp_t - \ln shhp_{t-1} \quad (6)$$

这里  $shhr_t$  表示上证综指第  $t$  期的日收益率,  $shhp_t$  表示上证综指第  $t$  期的日收盘价。图 1 分别描绘了上海股市日收盘价收益率的波动,波动具有随时间变化的特征,有时相当稳定,有时波动异常激烈,形成了收益率的波动聚集性,显示了收益率的时变方差特征。

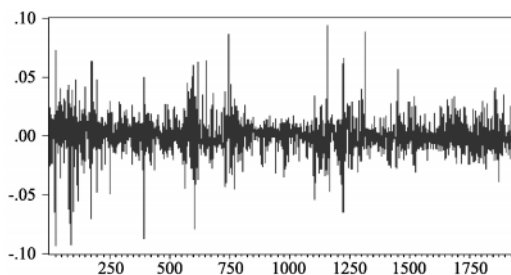


图 1 上海股市收益率波动图

为了取得波动率模型的最优样本容量,本文采用递归估计的方法,即以 1997 年 1 月 2 日至 1997 年 12 月 31 日,共一年的数据为初始样本,保持样本的起点不变,终点逐步增加一天交易日至样本中,当增加一个交易日到样本中,记样本为  $S_1$ ,样本容量为 244;第二个样本在第一个样本的基础上增加一个数据,记为  $S_2$ ,样本容量 245;以此类推第  $S_{1685}$  个样本的时间从 1997 年 1 月 2 日至 2004 年 12 月 31 日,样本容量 1928。在我们的实证分析中,通过建立不同阶数的模

型并进行比较,发现 AR(1)-GARCH(1,1)能够较好地拟合上海股市的波动。并且 Yew, McAleer 和 Ling (2002)认为在大量的金融时间序列实证分析中,AR(1)-GARCH(1,1)是比较流行使用的一种形式。因此,在所建立的 1685 个模型中,我们采用 AR(1)-GARCH(1,1)的形式。

#### 3.2 上海股市的最优样本容量

对于最优样本容量的确定,本文通过编程分别对样本  $S_1, S_2, \dots, S_{1685}$  建模,通过 1685 次递归估计,共建立 1685 个 GARCH 模型。图 2 为 GARCH 模型的估计参数,揭示了上海股市的  $\alpha$ (ARCH 项)和  $\beta$ (GARCH 项)参数  $\alpha$  和  $\beta$  在向着相反方向移动,当  $\alpha$  比较大时, $\beta$  值比较小,反之亦然。各个阶段参数稳定性的不同主要体现在三到四年(1999 年 1 月至 1999 年 12 月)和七到八年(2003 年 1 月至 2003 年 12 月)的样本容量,这两个期间模型的估计系数不同于其他阶段, $\alpha$  和  $\beta$  的系数比较稳定,没有产生剧烈的上下波动,而在其它期间系数常常会出现突然的跳跃或者大幅度波动,比如当选取样本的终点出现在两到三年(1998 年 1 月至 1998 年 12 月)时,即样本容量是两到三年, $\alpha$  系数会出现向上跳跃, $\beta$  系数会突然向下跳跃。

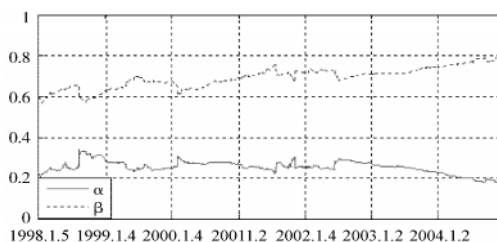


图 2 波动率 GARCH 估计参数递归图

为了反映各年之间的差异,分别计算各年  $\alpha$  和  $\beta$  的极差和标准差。极差是期间最大值与最小值的距离,反映参数的波动范围,标准差则说明参数的整体波动情况,二者结合可以说明参数的稳健性。从表 1 也可以看出,在三到四年和七到八年的期间,反映系数变动情况的极差和标准差都要小于相邻年份,从而上海股市 GARCH 模型的最优样本容量会出现在这两

表 1 对称 GARCH 模型系数的稳健性

系数	数据	年份						
		1998	1999	2000	2001	2002	2003	2004
$\alpha$	极差	0.1328	0.0588	0.0661	0.0847	0.0535	0.0477	0.0493
	标准差	0.0379	0.0150	0.0194	0.0133	0.0168	0.0120	0.0147
$\beta$	极差	0.0848	0.0690	0.0803	0.0911	0.0520	0.0401	0.0463
	标准差	0.0247	0.0187	0.0203	0.0149	0.0162	0.0114	0.0139

个阶段。从计算稳健的  $t$  统计量(图 3)知道,这两个期间的  $\alpha$  和  $\beta$  都是显著的。此外, GARCH(1,1) 的系数反映了条件峰度的水平,  $\alpha$  值越大(小),说明峰度越高(小)(Yew 等, 2002),这是由于极端值的出现增大了  $\alpha$  的值,即增大了条件峰度。

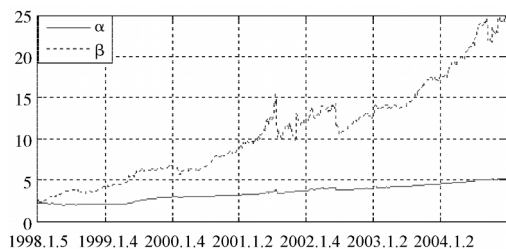


图 3 GARCH 估计参数  $t$  统计量递归图

为了有效地判断样本容量的稳健性,需要进一步考察衡量模型稳定的二阶矩条件和四阶矩条件(图 4),上海股市在整个时间周期展示了令人满意的递归二阶矩。对于四阶矩条件, GARCH 模型在七到八年的时间里,违反了模型的四阶稳定性条件,说明七到八年的样本容量不是合适的,而三到四年的样本容量,四阶矩得到令人满意的结果。要注意的是,虽然这一期间偶尔不满足四阶矩条件,这是由于上海股市是新生股票市场,波动幅度要大于如美国、日本和新加坡等成熟的股票市场,会出现一些极端值,从而造成四阶稳定性被违反。就整个时间期间而言,上海股市波动率建模稳健估计的样本容量是三到四年,这和 Yew 等(2002)使用香港恒生指数、日经 225 指数和 S&P500 指数对香港股市、日本股市和美国股市分析所得到的结论相同。

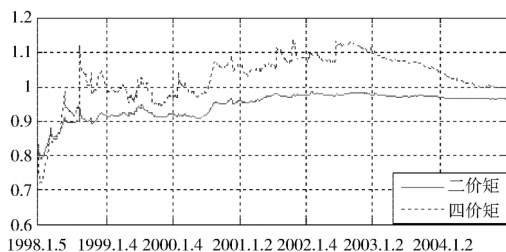


图 4 GARCH 模型二阶矩和四阶矩的递归图

## 4 思考

本文通过 GARCH 模型探讨了最优样本容量的确定问题,我们可能从以下两个方面考虑实证金融分析中如何确定最优样本容量:

(1) 根据中国金融制度的变化选择样本。中国经济处于转轨时期,经济结构的变化会导致不同时期的数据具有不同的特征,因此我们不宜跨不同时期选择样本。比如从 1993 年 1 月 1 日至 2007 年 12 月 31 日,中国股市共经历了两次交易制度的变化,可以分为 T+0 时期、T+1 时期以及有 10% 涨跌停板限制的 T+1 时期,而研究表明(陈珍珍、赵华, 2003),在进行实证分析时,必须考虑到交易制度的变化对股市运行的影响,不宜跨时期选择数据,如果忽略了这一点,简单地把三个时期放在一起综合分析,可能导致分析结论出现偏误。

(2) 分析不同样本容量对研究结论的影响。对于特定的实证研究,可能由于一些原因没有办法找出最优样本容量,这时我们通过尝试使用不同样本容量的样本进行分析,检验样本容量对参数的影响。如果参数对样本容量的变化比较敏感,说明选择样本容量必须慎重,以免得出有误的结论。

### 参 考 文 献

- [1] 陈珍珍、赵华:《不同交易制度下 CAPM 统计检验》,《厦门大学学报(哲社版)》2003 年第 3 期。
- [2] Bollerslev, T., "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 1986(31), 307-327.
- [3] Engle, R. F., "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, 1982(50), 987-1007.
- [4] Ling, S. and M. McAleer, "Stationarity and the existence of moments of a family of GARCH processes," *Journal of Econometrics*, 2002(106), 109-117.
- [5] Yew, X., M. McAleer, and S. Ling, "Determining an optimal window size for modeling volatility," *Handbook of Applied Econometrics and Statistical Inference*, 443-467, Marcel Dekker, Inc. 2002.

(责任编辑: 扈媛媛)